EXAMINING THE VALIDITY OF A GENERATIVE EDUCATION PATTERN BASED QUESTION

Karen Leary Duseau
University of Massachusetts, Dartmouth
kduseau@umassd.edu

Assessment is a topic of concern to all stakeholders in our educational system. Pattern Based Questions are an assessment tool which is an alternative to the standardized assessment tool, and they are based on generative learning pedagogy, which shows promise in engaging all learners and usefulness in teaching and learning but validity has not yet been empirically established. Pattern-based questions seek to provide a qualitative expression of student understanding. It is the purpose of this research to empirically explore a correspondence between student response patterns and the students' expressed ways of thinking using a grounded theory approach and clinical interviews. Findings include rich descriptions of participants' ways of thinking about equivalent fractions as written in the PBQ. One response pattern was clearly differentiated from the others. Future research is discussed.

Keywords: Assessment, Rational Numbers & Proportional Reasoning, Undergraduate Education.

Introduction

Assessment is a concern for all stakeholders in education. Current quantitative standardized assessment has been critiqued for focusing on procedural understanding instead of conceptual understanding (Herman et al., 1992; Kohn, 2000; Schwartz, 1991; Stake, 1995), for promoting inequities in education (Knoester & Au, 2017; Noble et al., 2012; Stroup, 2009), for not accurately representing student knowledge (Noble et al., 2012; Stake, 1995), and for promoting teaching to the tests (Riffert, 2005). Pattern-based questions (PBQs) are a qualitative alternative which can be used at scale. PBQs are an assessment tool associated with a larger pedagogical theory called generative education (Stroup et al., 2004; Wittrock, 1974). (in prior work PBQs have been referred to as PBIs – pattern-based items and NDMC – non dichotomous multiple choice.) Learning environments designed using generative education pedagogy can provide students with opportunities to engage in activities and gain a deep understanding of meaningful mathematics (Duseau, 2019). As a tool of generative education, PBQs promote engaging *all* students by utilizing a "low threshold, high ceiling" (Papert, 1980/1993) approach. Students may participate at a low level of understanding (low threshold) or at a high level of understanding (high ceiling). Pattern-based questions seek to provide a qualitative expression of student understanding which does not reduce student understanding down to a single number. Some Educators believe that the response patterns of PBQs identify patterns in student thinking. This allows students to show what they do know instead of focusing on their deficits. Educators have created PBQs using what they know about students' common understandings of mathematical material. To date, however, there has been no systematic investigation of the empirical warrant for the correspondence between students' response patterns on PBQs and students' verbalized understanding of mathematical concepts. Therefore, my research questions are:

- Can this pattern-based question identify students' ways of understanding equivalent fractions? If so, what are these ways of thinking?

- Can some correspondence be established between students' response patterns and students' verbalized understanding of mathematical concepts?

Theoretical Framework

PBQs are an assessment tool associated with the larger pedagogical theory of Generative Design. Generative design is a pedagogical theory of learning and teaching in which students generate artifacts in order to actively engage in the mathematics, make sense of the mathematics, make connections to other mathematics, and to be creative (Stroup et al., 2004). Stroup, Ares, and Hurford (2004) describe generative design as a "space creating" pedagogy in which students create "a space - a coordinated collection – of expressive artifacts and actions in relation to some shared task or set of rules" (p. 839). This means the students are creating artifacts such as equations, functions, or expressions which share some property (or rule) and the collection of those artifacts should enlighten the students to some structure in mathematics. PBQs can be developed using that same space, or collection of student created artifacts because they represent the students own conceptions of a mathematical idea. PBQs are intended to display the current state of students' understanding and support educators' analyses of those understandings.

This study framed the research through a constructivist lens and used grounded theory (Charmaz, 2014) and clinical interviews (Ginsburg, 1997) as tools. Charmaz's (2014) perspective on grounded theory is appropriate because she claims a constructivist view of grounded theory as opposed to Glaser and Strauss who have a more post-positivistic view. Generative design, as discussed previously and as linked to the development of PBQs, has emerged from the constructivist tradition.

Piaget is credited for introducing the clinical interview method to contrast with the standardized administration of tests. Piaget (1954) believed that children construct their own knowledge or understanding. His methodology of the clinical interview is aligned with his understanding of the constructivist perspective. Ginsburg (1997) wrote a more detailed description of the clinical interview and expounded upon the rationale for using the clinical interview. Ginsburg (1997) gave many reasons to use clinical interview methods, including 1) constructivist theory requires a method "that attempts to capture the distinctive nature of the child's thought" (p. 58), 2) clinical interviews allow us to deal with fluidity of thinking, 3) clinical interviews work for the individual or the general, and 4) clinical interviews "embody a special kind of methodological fairness, especially appropriate in a multicultural society" (p. 58). The clinical interview allowed us to use questioning to probe the understandings associated with participants' responses.

Methodology

For this study, participants were given an online assessment item via google forms. The online assessment was a "choose all" PBQ. After the assessment results were organized and examined using google sheets to organize and display response patterns, some participants were invited to participate in a clinical interview via Zoom. A retrospective protocol was implemented for these task-based interviews. The data was analyzed using Charmaz's (2014) three types of coding.

Participants

I solicited participants from a local State University in Massachusetts. This is an important population to study and understand because this large population of students are admitted into college and university and are not prepared to take college level courses in mathematics. The

current Massachusetts Curriculum Framework for Mathematics (DESE, 2017) proposes to prepare students to be college and career ready in mathematics. Specifically, they state "Students who are college and career ready in mathematics will at a minimum demonstrate the academic knowledge, skills, and practices necessary to enter into and succeed in entry-level, credit bearing courses in College Algebra, Introductory College Statistics, or technical courses" (DESE, 2017, p. 9). However, many students exit high school not yet prepared and ready to succeed in entry-level, credit bearing courses at the public college level. This is a current concern of many colleges, universities, and the students themselves.

The Assessment Item

The assessment item was selected for this proposed research for three reasons. First, this item has been refined and implemented at a statewide level prior to this research (Stroup, 2019). Second, the participants have been previously exposed to the fraction concepts. I chose fractional equivalence which appears much earlier in the typical mathematics sequence, to ensure that students would have experience with the topic and would be able to engage in the concepts. By the time students reach the undergraduate level, one would expect that they have years of experience with fraction concepts and will have some fairly stable understanding of those concepts. Lastly, the fractional equivalence represents important ideas that impact the students' success in their current courses. Fraction concepts continue to create obstacles to success for developmental mathematics students. Undergraduate developmental mathematics students have seen this material and this material still contributes to student difficulties in their current (and future) mathematics classes.

The assessment item discussed in this research is shown in Figure 1.

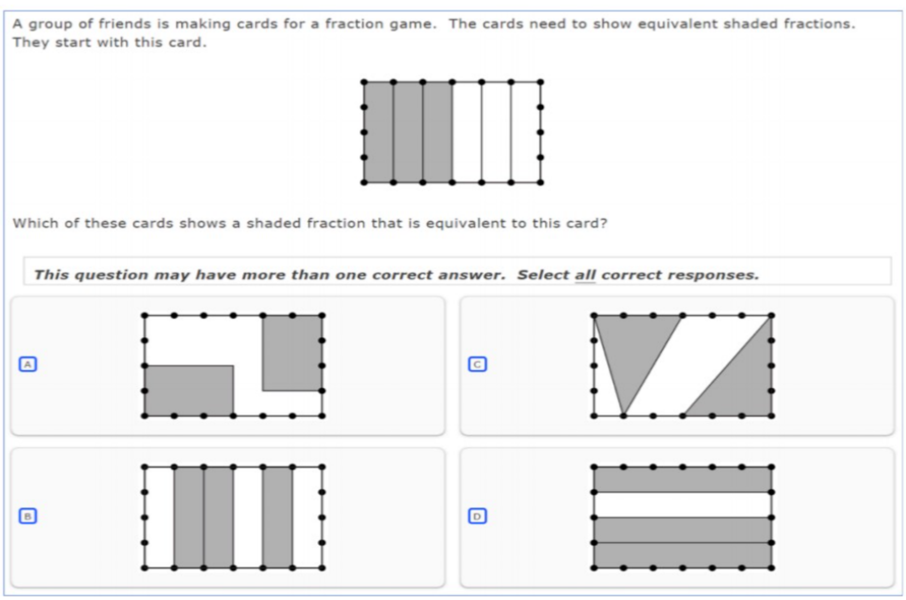

Figure 1: The assessment item, Equivalent Fractions

The "low threshold" option for this item is option B. Students need only a foundational understanding of fraction equivalence to identify that option B is equivalent to the original fraction card. Options A and C give students some steps towards the "high ceiling" in that a

deeper understanding of fractional equivalence is needed to understand that A and C are also equivalent to the original. Option D, which is not equivalent to the original card, introduces an alternative that can reveal important features of student thinking. For the response patterns of BD, teachers assume that the students are counting bars and for that way of thinking, any option that has three shaded bars will represent an equivalent fraction. This combination (BD) of one correct response and one incorrect response potentially gives us insight into how the student is thinking about the mathematics. This is an example of how a "correct" answer or an "incorrect" answer do not give the full picture, but the combination of response options makes clear the student thinking.

The Interview

Students with common response patterns were asked to participate in a virtual Zoom interview. The common response patterns will be informed by Stroup's (2019) *Interim Assessment Pattern-Based Item Report* and the response pattern diagram created in Google Sheets.

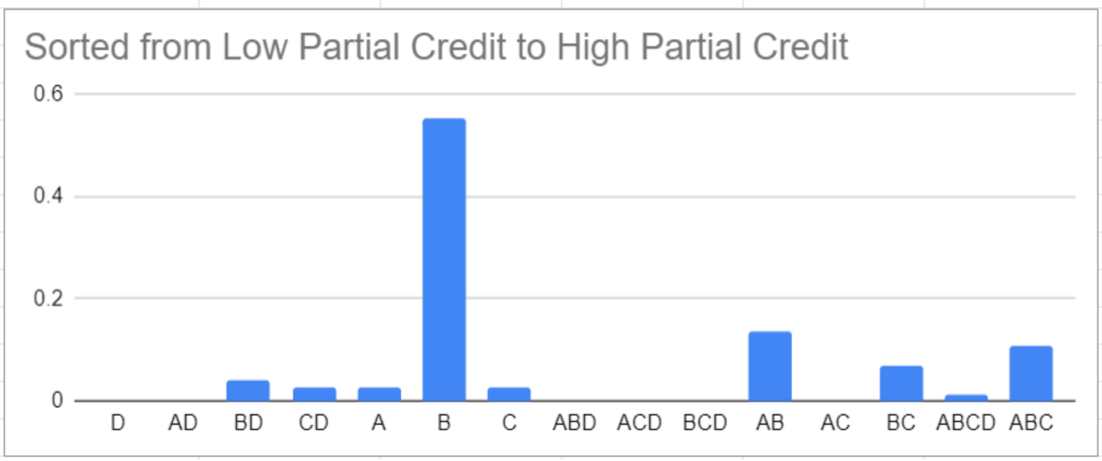

Figure 2: The Response Patterns Diagram from the Assessment Item

Common response patterns are seen as "bumps" or modes in the response pattern diagrams because they are response patterns which are selected by students more frequently than other possible response patterns. Specifically, I chose participants who chose the B combination because there were such a high percentage of participants who chose that option. Additionally, I chose participants who chose any combination include the D option to examine the "BD" type of thinking. I also chose an arbitrary number of other response patterns to open up possibilities I may not have foreseen. Unfortunately, not all students who were invited to participate in an interview were agreeable. I was not able to get as much "BD" data as I wanted to. Additionally, some of the students who began with the "BD" combination changed their answers once the interview began.

IRB consent forms were collected from participants before each interview. Interviews were conducted via Zoom to ensure safe COVID19 protocols. The zoom interview was recorded and automatically transcribed.

During the interview, I asked participants to walk me through their thinking, explain how they approached the question, or how they made the decision to choose their specific response pattern. I attempted to avoid the question "why" because this can be interpreted as a judgmental

question and I was attempting to establish a comfortable, “no judgement” zone atmosphere during the interviews. I also avoided any use of the words “correct” or “incorrect”. I was not interested in finding “right” or “wrong” answers, I was only looking for patterns in participants’ thinking.

Data Analysis

Data analysis was done using Charmaz’s (2014) three types of iterative coding and memo writing. Initial coding was done by using “in vivo” (Charmaz, 2014) codes which are special words or terms used by the participants that appeared to represent a way of thinking used by the participant. Focused coding examined the initial codes (and the text data) to see if two codes could be combined or if one code could subsume a second code. Charmaz (2014) tells us to “use focused coding to pinpoint and develop the most salient codes and then put them to the test with large batches of data” (p. 114). As the focused codes were developed the text data was referred to constantly. The codes were examined and considered iteratively and at length. Then the text data was again coded using the focused codes to ensure that the “ways of thinking” identified real patterns in participants thinking and the focus codes were not an artificial creation of manipulating the initial codes. This iterative step took quite back and forth work because in some instances, participants used similar words when they had different ways of thinking and in other instances participants used different words to describe similar ways of thinking. The participants had difficulties explaining their thinking, which made analyzing the text data challenging.

Lastly, theoretical coding can adopt concepts from the literature on the mathematical concepts to help structure emerging theories. For this research, there was limited structure. The concepts from the mathematics education literature will be compared to the ways of thinking (identified in the findings) in the discussion section.

Memo writing was used throughout the analysis process to identify and record anything which stood out to me or to identify ideas that emerged during the entire process. Most of the memo writing reflected participants affect, or feelings about doing and discussing mathematical thinking and does not directly reflect students’ ways of thinking.

Findings

Students’ ways of thinking about this question varied considerably. I have identified six different ways of thinking about fractional equivalence. I have restricted myself to ways of thinking that the participants gave explicit verbal evidence of. If they did not say it, I did not assume it. The following list describes the participants’ ways of thinking.

1. Visualize – Most students visually inspected the cards and tried to “match” to the original. Sometimes this involved “moving” and “fitting” pieces together. Some students were unable to visualize moving and fitting pieces together but could identify that “three rectangles shaded and three rectangles not shaded”, regardless of order, were equivalent.
2. “3 shaded and 3 not shaded” is different than “3 shaded and 1 not shaded”. After visually matching up the original with response selection B, I often asked about response selection D because it also had three shaded rectangles. Many of the students were able to differentiate between the two cards saying that three shaded with three unshaded was different than three shaded with one unshaded. In some cases, this may have been an indication of understanding a fractional ratio, but some participants were unable to verbalize this. In one case, the participant simply noticed that the original was “50 – 50”,

and three shaded with one not shaded was "more shaded than not" and so it is not equivalent to the original.

- ¾ not ½ - A number of students mentioned that three shaded with one unshaded was ¾ and that the original was ½ and that those were not the same. This might indicate the understanding of the fractional shading.
- Counting pieces – there were two different types of counting pieces.

1. One instance (Interview 6) of counting pieces was counting the number of shaded rectangles and matching three shaded rectangles to three shaded rectangles.
2. In two instances (Interview 4 and 8) the participants tried to count pieces of different shapes and sizes and force the shaded number of pieces into a fractional ratio with the total number of pieces.

- Rectangles only – In two instances (Interview 7 and 9) the participants considered only responses which were rectangles. They discounted response options A and C because there were shapes other than rectangles involved.
- Area – One participant (Interview 10) calculated the area of the shaded region of the original and compared that to the calculated area of each of the options. This participant missed option A because of a small mathematical error.

The following table indicates which interviews provided evidence of each type of reasoning. Additionally, this table gives their final response pattern and types of changes, if any.

Table 1: Summary Table of Findings. This table indicates the interview number, the response pattern, the types of reasoning evidenced, and the type of change, if any.

Int #	Res p.	Vis ual	Sha de vs not	¾ not ½	Cou nt pieces	Rec t only	Are a	Change
1	AB	*	*					Stable
2	BC	*		*				Stable
3	AB	*	*					Learni
4	B	*	*		*			Unstab
5	B	*	*					Stable
6	BD				*			Stable
7	B	*	*			*		Stable
8	B	*	*	*	*			Stable
9	B	*	*			*		Stable
10	BC	*					*	Learni

Changing Answers

Half the participants changed their responses during the interview. From my earliest memoing, students' changing their answers was a salient behavior, and it seemed to be relevant to the validity of these types of questions. The PBQ questions allowed students to modify their ways of thinking. I believe educators would be interested in knowing if a student was committed

to a solution or changed their answer and in what ways they may have made changes. All who changed answers moved from exhibiting a less "thought through" reasoning in the original response pattern to a more complete expression of understanding. There appear to be three types of changes involved. First is a stable change. Interview 1 changed from CD to ABC and interview 2 changed from AB to BC. Both of these changes were made immediately, and they could not describe why they had chosen the original response pattern. This scenario presents as though the participant was not initially motivated to think carefully about their response pattern. Once the change was made, the participants were stable in their thinking. The second type of change was called "active learning". Interview 3 changed from B to ABC and Interview 10 changed from C to BC. In these instances, participants made verbal indications that they were thinking about the questions as they were explaining their responses to me. Interview 3 made the comment "wait a minute, let me check …." and Interview 10 used the phrase "it's going to be like, wait, 1234, wait … Oh". These utterances seem to imply the participant was activing thinking about the problem. I called the last type of change unstable because the participant changed repeatedly and did not appear to understand why they were changing. Interview 4 changed from CD to B and used the phrases "It kind of would look the same, but, …", "honesty, I don't even know", "I think it might be B", and "I didn't think so, but now I feel like it does". These utterances do not indicate stability in their ways of thinking.

Discussion

The participants' ways of thinking about equivalent fractions is consistent with the literature on rational number concepts. The participants changing answers is a healthy part of learning and does not challenge the validity in respect to the goals of a qualitative PBQ. Correspondence between the BD response pattern and the "counting pieces" thinking was evidenced and the correspondence between the other response patterns and participants ways of thinking were a bit blurred.

Ways of Thinking and Rational Number Literature

The participants use of the "visualize" way of thinking where the participant matches up the three shaded rectangles and three not shaded rectangles is consistent with the literature on "part – whole" conception of fractions (Behr et al., 1983; Kieren, 1980; Lamon, 2007; Pederson & Bjerre, 2021). This conception frames the fraction equivalence as "The part-whole interpretation of rational number depends directly on the ability to partition either a continuous quantity or a set of discrete objects into equal-sized subparts or sets" (Behr et al., 1983, p. 93). This is seen clearly when the participants visualize three shaded and three not shaded rectangles. Furthermore, Behr et al. (1983) also refer to the part-whole conception as geometric regions and the participants certainly looked at the rectangles as geometric shapes.

The "three shaded and three not is different from three shaded and one not" way of thinking about fractional equivalence is consistent with ratio as a relationship between two quantities (Behr et al., 1983; Keiren, 1980; Lamon, 2007; Pederson & Bjerre, 2021) because the participants were able to see that the relationship between the number of shaded rectangles and the number of not shaded rectangles was different and therefor the fractions were not equivalent. What is not clear from the participants responses, and perhaps because of the assessment item itself is if the participants can differentiate the equivalence of "relative magnitudes" (Pedersen & Bjerre, 2021, p. 144) because all the cards in the assessment item were the same size.

The way of thinking which I called "¾ is not equal to ½" is consistent with the rational number concept of "fraction as a number" (Lamon, 2007, p. 635) because the participants were

able to name three shaded rectangles and three not shaded rectangles as the number ½ and they named three shaded rectangles and one not shaded rectangles as ¾ . Additionally, they understood that those two numbers were not an absolute equal amount. We cannot be sure if they understood if ¾ and ½ were not equal relative amounts.

The "calculate area" way of thinking is consistent with Behr et al.'s (1980) reference to geometric regions. Using the dots on the cards as unit lengths, this participant found the geometrical area of the original card and the four solution option cards and compared the absolute area to the absolute area.

"Counting pieces" and "Only rectangles" are limited conceptions of fractional equivalence. In both cases these ways of thinking show participants difficulties with rational number concepts. These ways of thinking are still very real to the participants and important for the educator to be aware of to help students grow and develop better understandings of rational number concepts.

Changing Answers

Changing one's response pattern is a natural extension of learning and as such it is encouraging that PBQs can have a role in providing a pathway, or context, for students to extend their understandings. There are a number of possible reasons for shifts in students' understandings. Two shifts discussed were the stable change when a student changes immediately because they were not motivated to think carefully about the item when they first attempted it, but are now more motivated because they need to explain their thinking to the interviewer. This also lends credence to the idea that explaining one's thinking is valuable both to the student to organize their thoughts but also to the assessor to have a clearer idea of what the students' thinking is. Another type of change was called "active learning". In this change, students were willing to think while explaining and the student had an "ah-ha" moment during the interview. Both forms of changing answers are seen as corresponding to expressed shifts, or clarification, in students understanding. These shifts would not make the items invalid because they no longer represent the students' thinking, validity is not fixed in relation to the initial responses but instead can include changes or updates that are what come to be seen as linked to selection a new combination of responses. Students whose understanding of a problem shifts from one way of thinking to another way of thinking can still find a response pattern which can represent the new way of thinking. Education involves developing more-fully integrated and robust forms of understanding. Assessment should be able to accommodate these shifts in understanding.

Correspondence between Response Patterns and Ways of thinking

There was a clear correspondence between the response pattern of BD and the counting shaded rectangles. The other response patterns of B, BC, and ABC were not clearly aligned or corresponding with one particular way of thinking. I believe this is because there is overlap between the different types of thinking displayed by "visualization", "3 shaded and 3 not shaded", and "¾ is not ½". These other types of thinking seem to stem from the part-whole conception. In a number of instances, the only difference between participants response patterns came down to small differences such as the ability to visualize if the shapes could "fit" together to create the same shaded shape as in the original card.

Conclusion

This PBQ does identify students' ways of thinking and in at least one instance there was a clear correspondence between the response pattern and the ways of thinking about fractional equivalence. Ideally, every PBQ would be researched and refined to empirically establish

validity. With validity established, PBQs could operate as an alternative to standardized assessments. Research should also be continued in the area of marginalized populations to establish if PBQs, which provide an opportunity for equity, can empirically be shown to provide a more equitable assessment tool.

References

Behr, M. J., Lesh, R., Post, T., & Silver, E. A. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 91 – 125). New York: Academic Press.

Charmaz, K. (2014). *Constructing grounded theory* (2nd edition). Sage Publications.

Duseau, K. (2019, April). Generative Design: The Teachers' Experience. [Paper Presentation]. The annual conference of the American Educational Research Association (AERA), Toronto, Canada.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press.

Herman, J. L., Aschbacher, P. R., Winters, L. (1992). *A practical guide to alternative assessment*. Association for Supervision and Curriculum Development, Alexandria, VA.

Kieren, T. E. (1980). The rational number construct: Its elements and mechanisms. In T. E. Kieren (Ed.), *Recent research on number learning* (pp 125 – 150). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Knoester, M., & Au, W. (2017). Standardized testing and school segregation: like tinder for fire? *Race Ethnicity and Education*, *20*(1), 1-14.

Kohn, A. (2000). Standardized testing and its victims. *Education Week*, *20*(4), 46-47.

Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework for research. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629-667). Reston, VA: Lawrence Erlbaum.

Massachusetts Department of Elementary & Secondary Education (DESE), (2017). *Massachusetts Curriculum Framework – 2017: Mathematics grades Pre-Kindergarten to 12*, Massachusetts Department of Elementary & Secondary Education. https://www.doe.mass.edu/frameworks/math/2017-06.pdf

Noble, T., Suarez, C., Rosebery, A., O'Connor, M. C., Warren, B., & Hudicourt-Barnes, J. (2012). "I never thought of it as freezing": How students answer questions on large-scale science tests and what they know about science. Journal of Research in Science Teaching, 49(6), 778-803.

Papert, S. A. (1993). Mindstorms: Children, computers, and powerful ideas, 2nd edition. Basic books. (Original work published in 1980)

Pedersen, P. L., & Bjerre, M. (2021). Two conceptions of fraction equivalence. Educational Studies in Mathematics, 107(1), 135-157. https://doi.org/10.1007/s10649-021-10030-7

Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). Basic Book, Inc. publishers, New York.

Riffert, F. (2005). The use and misuse of standardized testing: A Whiteheadian point of view. *Interchange*, *36*(1-2), 231-252.

Schwartz, J. L. (1991). The intellectual costs of secrecy in mathematics assessment. In V. Perrone (Ed.), *Expanding student assessment*, (pp. 132-141). The Association for Supervision and Curriculum Development (ASCD).

Stake, R. E. (1995). The invalidity of standardized testing for measuring mathematics achievement. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment*. (pp. 173 – 235). State University of New York Press.

Stroup, W. M., Ares, N. M., & Hurford, A. (2004). A taxonomy of generative activity design supported by next-generation classroom networks. In Proceedings of the twenty-sixth annual meeting of the Psychology of Mathematics Education, (pp. 1401-1410). Toronto: OISE/UT.

Stroup, W. (2009). What Bernie Madoff can teach us about accountability in education. Education Weekly, 28(25), 22-23.

Stroup, W. M. (2019, October). Interim Assessment Pattern-Based Item Report. GenEd Corporation.

Wittrock, M. C. (1974). A generative model of mathematics learning. *Journal for Research in Mathematics Education*, 181-196.